

The Evaluation Campaigns: Past, Present and Future

(Invited Talk)

Donna Harman

donna.harman@nist.gov

Scientist Emeritus, National Institute of Standards and Technology, USA

1 ABSTRACT

Evaluation has always been a critical component of information retrieval, and there have been some type of shared evaluations since the 1960s. The Cranfield test collection was used by multiple groups, starting with Gerard Salton in the 60s, and then by researchers at the University of Cambridge during the 1970s. But different versions of the collection were used and there was little attempt to compare results across systems. The creation of the large TIPSTER collection in 1990, followed by the first Text REtrieval Conference (TREC) in 1992 reframed the shared concept to mean not only using the same test collection, but also having a specific shared task, which in 1992 was an adhoc search tasks for 50 topics. Researchers could compare systems, and then incorporate what was jointly learned into their own systems. This paradigm grew in TREC to encompass new community information retrieval tasks, such as question answering and working with web data. It also branched into other new areas, such as video retrieval (which was spun off into TRECvid), and cross-language retrieval, which led to the formation of the European CLEF in 2000. Other shared evaluations like NTCIR in Japan and FIRE in India were organized, each targeting retrieval tasks most pertinent to their research communities. All of these evaluations have evolved over the years as the interests of the research groups have changed,

with evaluations in 2019 tackling problems of tracking emergency situations by following tweet streams, identifying birds by their calls, or working with Lifelogs.

2 BIOGRAPHY

Donna Harman graduated from Cornell University with a degree in electrical engineering, and having worked with Professor Gerard Salton, has been involved with research in new search engine techniques for many years. She retired from the National Institute of Standards and Technology in 2005 after leading a group that worked in the area of natural language access to full text. In 1992 she started the Text REtrieval Conference (TREC), a still-ongoing forum that brings together researchers from industry and academia to test their search engines against common corpora. She received the 1999 Strix Award from the UK Institute of Information Scientists for this effort. She is currently a scientist emeritus at NIST and is the author of two textbooks: *Information Retrieval Evaluation* and a new history book, *Information Retrieval: the Early Years*.

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

IIR 2019, September 16–18, 2019, Padova, Italy